

0400  
04-30-01 0280

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/819,136

DATE: 04/10/2001

TIME: 15:49:44

Input Set : A:\0025.txt

Output Set: N:\CRF3\04102001\I819136.raw

4 <110> APPLICANT: Conklin, Darrell C.  
5 Gao, Zeren  
7 <120> TITLE OF INVENTION: MULTI-DOMAIN PROTEINASE INHIBITOR  
9 <130> FILE REFERENCE: 00-25  
C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/819,136  
C--> 11 <141> CURRENT FILING DATE: 2001-03-27  
11 <150> PRIOR APPLICATION NUMBER: US 60/193,642  
12 <151> PRIOR FILING DATE: 2000-03-31  
14 <160> NUMBER OF SEQ ID NOS: 13  
16 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
18 <210> SEQ ID NO: 1  
19 <211> LENGTH: 2082  
20 <212> TYPE: DNA  
21 <213> ORGANISM: Homo sapiens  
23 <220> FEATURE:  
24 <221> NAME/KEY: CDS  
25 <222> LOCATION: (376)...(2022)  
27 <400> SEQUENCE: 1

28	gtgaccctca tggccaqtgg ctctgtgctc atgqqcctct gcccctccc caacctctc	60
29	ccctctgccc tgtgtgacc agggcctggg agcccccgca cgggtcagac agaggggccc	120
30	ggctgaagct ggagaggaac cagcgtcaca cagacggcct ctgagaactt ggagaccccq	180
31	ttaccacccc agcaggggtg tcaggacaag catctgtctg aggttcagc ctcaggggca	240
32	aaaggagacc ccgggtcct ggtgggggca ccgaccacag gcccgagggt tggatgctg	300
33	caggaagctg ggctctgtgg agcccgagga ggggtgggtg gccacacccc ccggccccct	360
34	ggctcggcgg cctc atg ccc gcc cta cgt cca ctc ctg ccg ctc ttg ctc	411
35	Met Pro Ala Leu Arg Pro Leu Leu Pro Leu Leu Leu	
36	1 5 10	
38	ctc ctc cgg ctg acc tgg ggg gct ggc ttg ctg cca ggg ctg ggg agc	459
39	Leu Leu Arg Leu Thr Ser Gly Ala Gly Leu Leu Pro Gly Leu Gly Ser	
40	15 20 25	
44	cac tcc ggc gtg tgc ccc aac cag ctc agc ccc aac ctg tgg gtg gac	507
45	His Pro Gly Val Cys Pro Asn Gln Leu Ser Pro Asn Leu Trp Val Asp	
46	30 35 40	
48	gcc cag agc acc tgt gag cgc gag tgt agc agg gac cag gac tgt gcg	555
49	Ala Gln Ser Thr Cys Glu Arg Glu Cys Ser Arg Asp Gln Asp Cys Ala	
50	45 50 55 60	
52	gct gct gag aag tgc tgc atc aac gtg tgt gga ctg cac agc tgc gtg	603
53	Ala Ala Glu Lys Cys Cys Ile Asn Val Cys Gly Leu His Ser Cys Val	
54	65 70 75	
56	gca gca cgc ttc ccc ggc agc cca gct qcg ccg acg aca gcg gcc tcc	651
57	Ala Ala Arg Phe Pro Gly Ser Pro Ala Ala Pro Thr Thr Ala Ala Ser	
58	80 85 90	
60	tgc gag ggc ttt gtg tgc cca cag cag ggc tgc gac tgc gac atc tgg	699
61	Cys Glu Gly Phe Val Cys Pro Gln Gln Gly Ser Asp Cys Asp Ile Trp	
62	95 100 105	
64	gac ggg cag ccc gtg tgc cgc tgc cgc gac cgc tgt gag aag gag ccc	747
65	Asp Gly Gln Pro Val Cys Arg Cys Arg Asp Arg Cys Glu Lys Glu Pro	

P-5  
ENTERED

## RAW SEQUENCE LISTING

DATE: 04/10/2001

PATENT APPLICATION: US/09/819,136

TIME: 15:49:44

Input Set : A:\0025.txt

Output Set: N:\CRF3\04102001\I819136.raw

66	110	115	120	
68	agc ttc acc tgc gcc tgc gac ggc ctc acc tac tac aac cgc tgc tat	795		
69	Ser Phe Thr Cys Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr			
70	125 130 135 140			
72	atg gac gcc gag gcc tgc ctg cgg ggc ctg cac ctc cac atc gtg ccc	843		
73	Met Asp Ala Glu Ala Cys Leu Arg Gly Leu His Leu His Ile Val Pro			
74	145 150 155			
76	tgc aag cac gtg ctc agc tgg ccg ccc agc agc ccg ggg ccg ccg gag	891		
77	Cys Lys His Val Leu Ser Trp Pro Pro Ser Ser Pro Gly Pro Pro Glu			
78	160 165 170			
80	acc act gcc cgc ccc aca cct ggg gcc gcg ccc gtg cct cct gcc ctg	939		
81	Thr Thr Ala Arg Pro Thr Pro Gly Ala Ala Pro Val Pro Pro Ala Leu			
82	175 180 185			
84	tac agc agc ccc tcc cca cag gcg gtg cag gtt ggg ggt acg gcc agc	987		
85	Tyr Ser Ser Pro Ser Pro Gln Ala Val Gln Val Gly Gly Thr Ala Ser			
86	190 195 200			
88	ctc cac tgc gac gtc agc ggc cgc ccg ccg cct gct gtg acc tgg gag	1035		
89	Leu His Cys Asp Val Ser Gly Arg Pro Pro Pro Ala Val Thr Trp Glu			
90	205 210 215 220			
92	aag cag agt cac cag cga gag aac ctg atc atg cgc cct gat cag atg	1083		
93	Lys Gln Ser His Gln Arg Glu Asn Leu Ile Met Arg Pro Asp Gln Met			
94	225 230 235			
96	tat ggc aac gtg gtg gtc acc agc atc ggg cag ctg gtg ctc tac aac	1131		
97	Tyr Gly Asn Val Val Val Thr Ser Ile Gly Gln Leu Val Leu Tyr Asn			
98	240 245 250			
100	gcg cgg ccc gaa gac gcc ggc ctg tac acc tgc acc gcg cgc aac gct	1179		
101	Ala Arg Pro Glu Asp Ala Gly Leu Tyr Thr Cys Thr Ala Arg Asn Ala			
102	255 260 265			
104	gct ggg ctg ctg cgg gct gac ttc cca ctc tct gtg gtc cag cga gag	1227		
105	Ala Gly Leu Leu Arg Ala Asp Phe Pro Leu Ser Val Val Gln Arg Glu			
106	270 275 280			
108	ccg gcc agg gac gca gcc ccc agc atc cca gcc ccg gcc gag tgc ctg	1275		
109	Pro Ala Arg Asp Ala Ala Pro Ser Ile Pro Ala Pro Ala Glu Cys Leu			
110	285 290 295 300			
112	ccg gat gtg cag gcc tgc acg ggc ccc act tcc cca cac ctt gtc ctc	1323		
113	Pro Asp Val Gln Ala Cys Thr Gly Pro Thr Ser Pro His Leu Val Leu			
114	305 310 315			
116	tgg cac tac gac ccg cag cgg ggc ggc tgc atg acc ttc ccg gcc cgt	1371		
117	Trp His Tyr Asp Pro Gln Arg Gly Gly Cys Met Thr Phe Pro Ala Arg			
118	320 325 330			
120	ggc tgt gat ggg gcg gcc cgc ggc ttt gag acc tac gag gca tgc cag	1419		
121	Gly Cys Asp Gly Ala Ala Arg Gly Phe Glu Thr Tyr Glu Ala Cys Gln			
122	335 340 345			
124	cag gcc tgt gcc cgc ggc ccc ggc gac gcc tgc gtg ctg cct gcc gtg	1467		
125	Gln Ala Cys Ala Arg Gly Pro Gly Asp Ala Cys Val Leu Pro Ala Val			
126	350 355 360			
130	cag ggc ccc tgc cgg ggc tgg gag ccg cgc tgg gcc tac agc ccg ctg	1515		
131	Gln Gly Pro Cys Arg Gly Trp Glu Pro Arg Trp Ala Tyr Ser Pro Leu			
132	365 370 375 380			

## RAW SEQUENCE LISTING

DATE: 04/10/2001

PATENT APPLICATION: US/09/819,136

TIME: 15:49:44

Input Set : A:\0025.txt

Output Set: N:\CRF3\04102001\I819136.raw

```

134 ctg cag cag tgc cat ccc ttc gtg tac ggt ggc tgc gag ggc aac ggc 1563
135 Leu Gln Gln Cys His Pro Phe Val Tyr Gly Gly Cys Glu Gly Asn Gly
136          385          390          395
138 aac aac ttc cac agc cgc gag agc tgc gag gat gcc tgc ccc gtg ccg 1611
139 Asn Asn Phe His Ser Arg Glu Ser Cys Glu Asp Ala Cys Pro Val Pro
140          400          405          410
142 cgc aca ccg ccc tgc cgc gcc tgc cgc ctc cgg agc aag ctg gcg ctg 1659
143 Arg Thr Pro Cys Arg Ala Cys Arg Leu Arg Ser Lys Leu Ala Leu
144          415          420          425
146 agc ctg tgc cgc agc gac ttc gcc atc gtg ggg cgg ctc acg gag gtg 1707
147 Ser Leu Cys Arg Ser Asp Phe Ala Ile Val Gly Arg Leu Thr Glu Val
148          430          435          440
150 ctg gag gag ccc gag gcc gcc ggc ggc atc gcc cgc gtg gcg ctc gag 1755
151 Leu Glu Glu Pro Glu Ala Ala Gly Gly Ile Ala Arg Val Ala Leu Glu
152          445          450          455          460
154 gac gtg ctc aag gat gac aag atg ggc ctc aag ttc ttg ggc acc aag 1803
155 Asp Val Leu Lys Asp Asp Lys Met Gly Leu Lys Phe Leu Gly Thr Lys
156          465          470          475
158 tac ctg gag gtg acg ctg agt ggc atg gac tgg gcc tgc ccc tgc ccc 1851
159 Tyr Leu Glu Val Thr Leu Ser Gly Met Asp Trp Ala Cys Pro Cys Pro
160          480          485          490
162 aac atg acg gcg ggc gac ggg ccg ctg gtc atc atg ggt gag gtg cgc 1899
163 Asn Met Thr Ala Gly Asp Gly Pro Leu Val Ile Met Gly Glu Val Arg
164          495          500          505
166 gat ggc gtg gcc gtg ctg gac gcc ggc aac tac gtc cgc gcc gcc agc 1947
167 Asp Gly Val Ala Val Leu Asp Ala Gly Ser Tyr Val Arg Ala Ala Ser
168          510          515          520
170 gag aag cgc gtc aag aag atc ttg gag ctg ctg gag aag cag gcc tgc 1995
171 Glu Lys Arg Val Lys Lys Ile Leu Glu Leu Leu Glu Lys Gln Ala Cys
172          525          530          535          540
174 gag ctg ctc aac cgc ttc cag gac tag ccccccagcagg ggccctgcgcc 2042
175 Glu Leu Leu Asn Arg Phe Gln Asp *
176          545
178 accccgtcct ggtgaataaaa cgcactccct gtgcctcaga 2082
180 <210> SEQ ID NO: 2
181 <211> LENGTH: 548
182 <212> TYPE: PRT
183 <213> ORGANISM: Homo sapiens
185 <400> SEQUENCE: 2
186 Met Pro Ala Leu Arg Pro Leu Leu Pro Leu Leu Leu Leu Arg Leu
187 1 5 10 15
188 Thr Ser Gly Ala Gly Leu Leu Pro Gly Leu Gly Ser His Pro Gly Val
189 20 25 30
190 Cys Pro Asn Gln Leu Ser Pro Asn Leu Trp Val Asp Ala Gln Ser Thr
191 35 40 45
192 Cys Glu Arg Glu Cys Ser Arg Asp Gln Asp Cys Ala Ala Ala Glu Lys
193 50 55 60
194 Cys Cys Ile Asn Val Cys Gly Leu His Ser Cys Val Ala Ala Arg Phe
195 65 70 75 80

```

## RAW SEQUENCE LISTING

DATE: 04/10/2001

PATENT APPLICATION: US/09/819,136

TIME: 15:49:44

Input Set : A:\0025.txt

Output Set: N:\CRF3\04102001\I819136.raw

```

196 Pro Gly Ser Pro Ala Ala Pro Thr Thr Ala Ala Ser Cys Glu Gly Phe
197      85      90      95
198 Val Cys Pro Gln Gln Gly Ser Asp Cys Asp Ile Trp Asp Gly Gln Pro
199      100     105     110
200 Val Cys Arg Cys Arg Asp Arg Cys Glu Lys Glu Pro Ser Phe Thr Cys
201      115     120     125
202 Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr Met Asp Ala Glu
203      130     135     140
204 Ala Cys Leu Arg Gly Leu His Leu His Ile Val Pro Cys Lys His Val
205      145     150     155     160
206 Leu Ser Trp Pro Pro Ser Ser Pro Gly Pro Pro Glu Thr Thr Ala Arg
207      165     170     175
208 Pro Thr Pro Gly Ala Ala Pro Val Pro Pro Ala Leu Tyr Ser Ser Pro
209      180     185     190
210 Ser Pro Gln Ala Val Gln Val Gly Thr Ala Ser Leu His Cys Asp
211      195     200     205
212 Val Ser Gly Arg Pro Pro Pro Ala Val Thr Trp Glu Lys Gln Ser His
213      210     215     220
214 Gln Arg Glu Asn Leu Ile Met Arg Pro Asp Gln Met Tyr Gly Asn Val
215      225     230     235     240
216 Val Val Thr Ser Ile Gly Gln Leu Val Leu Tyr Asn Ala Arg Pro Glu
217      245     250     255
218 Asp Ala Gly Leu Tyr Thr Cys Thr Ala Arg Asn Ala Ala Gly Leu Leu
219      260     265     270
220 Arg Ala Asp Phe Pro Leu Ser Val Val Gln Arg Glu Pro Ala Arg Asp
221      275     280     285
222 Ala Ala Pro Ser Ile Pro Ala Pro Ala Glu Cys Leu Pro Asp Val Cln
223      290     295     300
224 Ala Cys Thr Gly Pro Thr Ser Pro His Leu Val Leu Trp His Tyr Asp
225      305     310     315     320
226 Pro Gln Arg Gly Gly Cys Met Thr Phe Pro Ala Arg Gly Cys Asp Gly
227      325     330     335
228 Ala Ala Arg Gly Phe Glu Thr Tyr Glu Ala Cys Gln Gln Ala Cys Ala
229      340     345     350
230 Arg Gly Pro Gly Asp Ala Cys Val Leu Pro Ala Val Gln Gly Pro Cys
231      355     360     365
232 Arg Gly Trp Glu Pro Arg Trp Ala Tyr Ser Pro Leu Leu Gln Gln Cys
233      370     375     380
234 His Pro Phe Val Tyr Gly Gly Cys Glu Gly Asn Gly Asn Asn Phe His
235      385     390     395     400
236 Ser Arg Glu Ser Cys Glu Asp Ala Cys Pro Val Pro Arg Thr Pro Pro
237      405     410     415
238 Cys Arg Ala Cys Arg Leu Arg Ser Lys Leu Ala Leu Ser Leu Cys Arg
239      420     425     430
240 Ser Asp Phe Ala Ile Val Gly Arg Leu Thr Glu Val Leu Glu Glu Pro
241      435     440     445
242 Glu Ala Ala Gly Gly Ile Ala Arg Val Ala Leu Glu Asp Val Leu Lys
243      450     455     460
244 Asp Asp Lys Met Gly Leu Lys Phe Leu Gly Thr Lys Tyr Leu Glu Val

```

## RAW SEQUENCE LISTING

DATE: 04/10/2001

PATENT APPLICATION: US/09/819,136

TIME: 15:49:44

Input Set : A:\0025.txt

Output Set: N:\CRF3\04102001\I819136.raw

```

245 465          470          475          480
246 Thr Leu Ser Gly Met Asp Trp Ala Cys Pro Cys Pro Asn Met Thr Ala
247          485          490          495
248 Gly Asp Gly Pro Leu Val Ile Met Gly Glu Val Arg Asp Gly Val Ala
249          500          505          510
250 Val Leu Asp Ala Gly Ser Tyr Val Arg Ala Ala Ser Glu Lys Arg Val
251          515          520          525
252 Lys Lys Ile Leu Glu Leu Leu Lys Gln Ala Cys Glu Leu Leu Asn
253          530          535          540
254 Arg Phe Gln Asp
255 545
257 <210> SEQ ID NO: 3
258 <211> LENGTH: 55
259 <212> TYPE: PRT
260 <213> ORGANISM: Homo sapiens
262 <400> SEQUENCE: 3
263 Thr Asp Ile Cys Lys Leu Pro Lys Asp Glu Gly Thr Cys Arg Asp Phe
264 1          5          10          15
265 Ile Leu Lys Trp Tyr Tyr Asp Pro Asn Thr Lys Ser Cys Ala Arg Phe
266          20          25          30
267 Trp Tyr Gly Gly Cys Gly Gly Asn Glu Asn Lys Phe Gly Ser Gln Lys
268          35          40          45
269 Glu Cys Glu Lys Val Cys Ala
270          50          55
272 <210> SEQ ID NO: 4
273 <211> LENGTH: 6
274 <212> TYPE: PRT
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: Glu-Glu tag
280 <400> SEQUENCE: 4
281 Glu Tyr Met Pro Met Glu
282 1          5
284 <210> SEQ ID NO: 5
285 <211> LENGTH: 55
286 <212> TYPE: PRT
287 <213> ORGANISM: Artificial Sequence
289 <220> FEATURE:
290 <223> OTHER INFORMATION: peptide motif
292 <221> NAME/KEY: VARIANT
293 <222> LOCATION: (2)...(7)
294 <223> OTHER INFORMATION: Xaa = any amino acid
296 <221> NAME/KEY: VARIANT
297 <222> LOCATION: (8)...(9)
298 <223> OTHER INFORMATION: Xaa = any amino acid or is not present
300 <221> NAME/KEY: VARIANT
301 <222> LOCATION: (11)...(25)
302 <223> OTHER INFORMATION: Xaa = any amino acid
304 <221> NAME/KEY: VARIANT

```

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/819,136

DATE: 04/10/2001  
TIME: 15:49:45

Input Set : A:\0025.txt  
Output Set: N:\CRF3\04102001\I819136.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:313 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:315 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:317 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:347 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:348 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:349 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:350 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:351 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:352 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:353 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:354 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:355 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:356 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:357 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:358 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:360 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:361 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:363 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:364 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:366 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:367 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:368 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:369 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:370 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:371 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:372 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:374 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7